# ML Project Requirements - Employee Salary Prediction

## 1. Software Requirements

- Python 3.8 or higher  
- Jupyter Notebook / VS Code / PyCharm  
- Libraries: pandas, numpy, scikit-learn, matplotlib, seaborn, tkinter (optional), joblib/pickle

## 2. Install Required Libraries

pip install pandas numpy scikit-learn matplotlib seaborn

## 3. Dataset Requirement

Dataset should include features like:  
- Job Title (e.g., Software Engineer)  
- Experience (e.g., 3, 5, 10 years)  
- Education Level (e.g., Bachelor's, Master's, PhD)  
- Location (e.g., New York, Bangalore)  
- Company Size (e.g., Small, Medium, Large)  
- Remote Work (Yes/No)  
- Salary (Target Variable)  
  
Sources: Kaggle datasets like "Salary Prediction Dataset", "Glassdoor Salary Data", etc.

## 4. ML Models & Techniques

Preprocessing:  
- Handle missing values: SimpleImputer  
- Encode categorical features: OneHotEncoder  
- Scale features: StandardScaler  
  
Models:  
- RandomForestRegressor  
- GradientBoostingRegressor  
- VotingRegressor (ensemble of above models)  
  
Evaluation Metrics:  
- MAE (Mean Absolute Error)  
- R² Score

## 5. Project Pipeline

1. Load and clean data  
2. Split into train/test sets  
3. Preprocess with ColumnTransformer  
4. Train individual models  
5. Combine using VotingRegressor  
6. Evaluate performance  
7. Optional: Build GUI or deploy web app

## 6. Optional GUI / Deployment

- Tkinter: Desktop GUI  
- Flask/Streamlit: Web app interface  
- Colab/Heroku/AWS: Deployment platforms